

FASTMath: Frameworks, Algorithms and Scalable Technologies for Mathematics

Lori Diachin, LLNL Institute Director



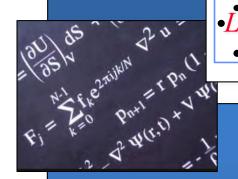
The FASTMath project brings leading edge computational mathematics technologies to the SciDAC Program

Develop advanced numerical techniques for DOE applications

- Eight focused topical areas based on application needs
- High level synergistic techniques

Deploy high-performance software on DOE supercomputers

- Algorithmic and implementation scalability
- Performance Portability
- Interoperability of libraries





FASTMath Objective:

Reduce the barriers facing application scientists



Demonstrate basic research technologies from applied mathematics

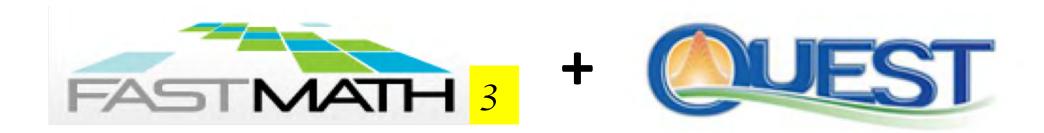
- Build from existing connections with basic research
- Focus on research results that are most likely to meet application needs

Engage and support of the computational science community

- Publications and presentations in highly visible venues
- Team tutorials
- Workforce pipeline and training
- Web presence



The SciDAC-4 FASTMath Institute leverages and builds on the successes of SciDAC-3 to meet application needs



- + Numerical Optimization
- + Data Analytics



FASTMath brings together an exceptional team of researchers and software library capabilities

Our team comprises over 50 researchers from 5 national laboratories and 5 universities



















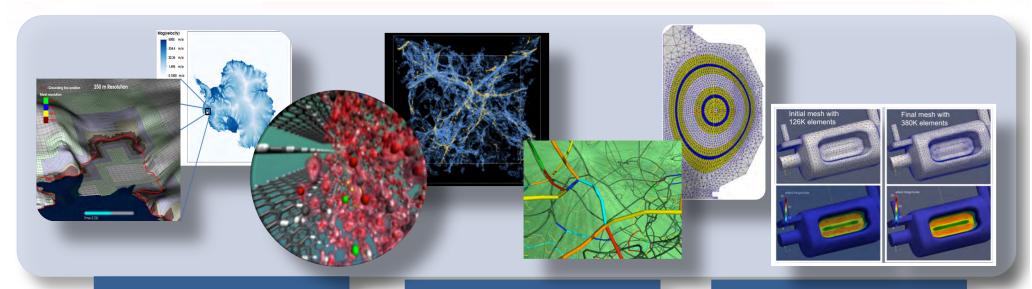








The FASTMath team has a proven record of advancing application simulation codes



Next Generation Application Codes

- Created unique DOE capabilities in ice sheet modeling
- First ever, self consistent solution of continuum gyrokinetic system in edge plasmas
- Unprecedented resolution for Nyx cosmology code

Faster Time to Solution

- New eigensolvers 2X faster for quantum chemistry software
- Parallel PEXSI software enabled electronic structure calculations with 10,000 atoms (compared to 1000's)
- Accelerated nonlinear solver enabled largest dislocation dynamics simulation with ParaDiS

More Robust Simulations

- Dramatically decreased time to generate meshes for fusion tokamak codes
- Adaptive mesh refinement and discretizations to resolve ELM disruptions in tokamaks
- Order of magnitude improvement in accuracy of integral calculations in material chemistry



For more information contact: Lori Diachin, LLNL diachin2@llnl.gov



FASTMath is actively engaged with 19 SciDAC-4 application partnerships

BER (5)

- Structured grid AMR
- Unstructured grid AMR
- Time integration
- Linear/Nonlinear solvers,
 Preconditioners
- Optimization
- Verification / UQ

NP (2)

- Structured grid AMR
- Eigenvalue problems
- Inference and Machine Learning

FES (5)

- Unstructured meshes
- Discretization technologies
- Iterative Linear solvers
- UQ

HEP (3)

- Direct solvers
- Structured Grid AMR
- Optimization
- Sensitivity Analysis
- Inference and machine learning

BES (2)

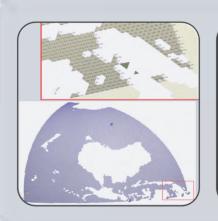
- Nonlinear and tensor eigenvalue problems
- Linear solvers and Preconditioners

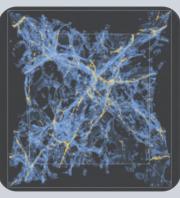
NE (1)

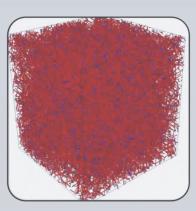
• UQ

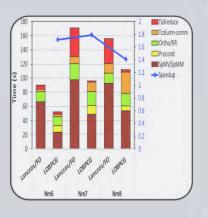


Math-CS collaborations led to significant advances in SciDAC-3 application sciences









Increased
accuracy of MPAS
Ocean modeling
system through
improved scaling
and improved
partitioning
algorithms

Development of highly-scalable, many core-aware multigrid solver increased ability of cosmologies to simulated evolution and fate of the universe (HPGMG in BoxLib)

New nonlinear solver and evaluation of OpenMP performance enabled largest dislocation dynamics simulation using ParaDiS

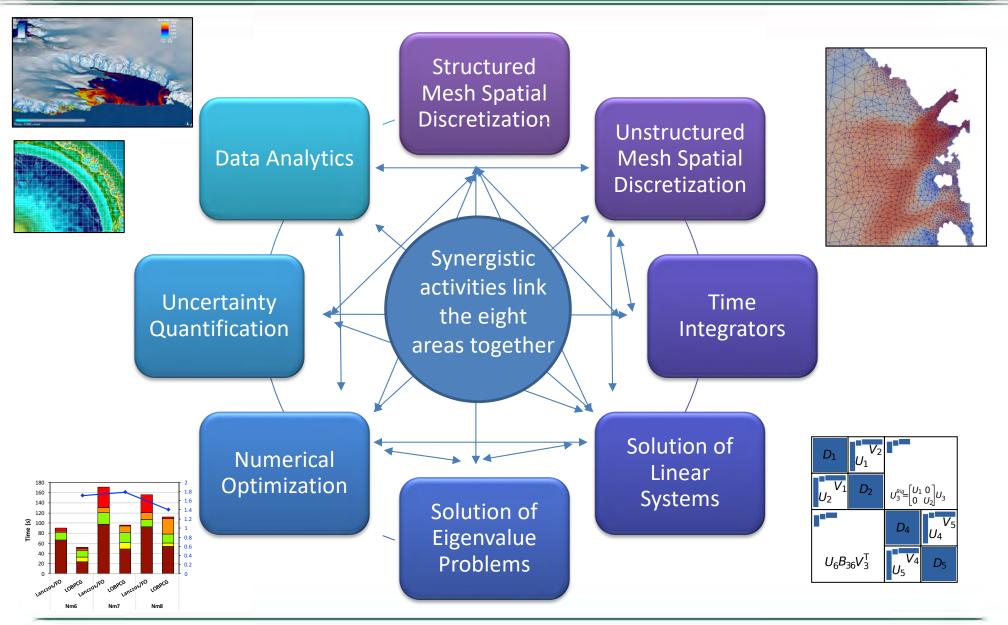
Improved
efficiency of
eigensystem
solution reduced
time to solution
in large, sparse ab
initio calculations
in nuclear
structure

FASTMath and RAPIDS will actively collaborate to continue to improve math libraries and application experience

- Performance improvements to math library software
 - Improved scaling (identify performance bottlenecks, find 'performance bugs', eliminate unnecessary communication)
 - Improved on-node performance (programming models, memory)
- Using performance models to improve foundational understanding of algorithms
- Advanced visualization tools for FASTMath tools (e.g., AMR)
- In situ visualization tools used in unstructured mesh simulation workflow
- Use of CS abstractions to improve or accelerate application development
 - Domain Specific Language compilers/tools
 - Leverage abstractions developed by RAPIDS for I/O to unify application experience



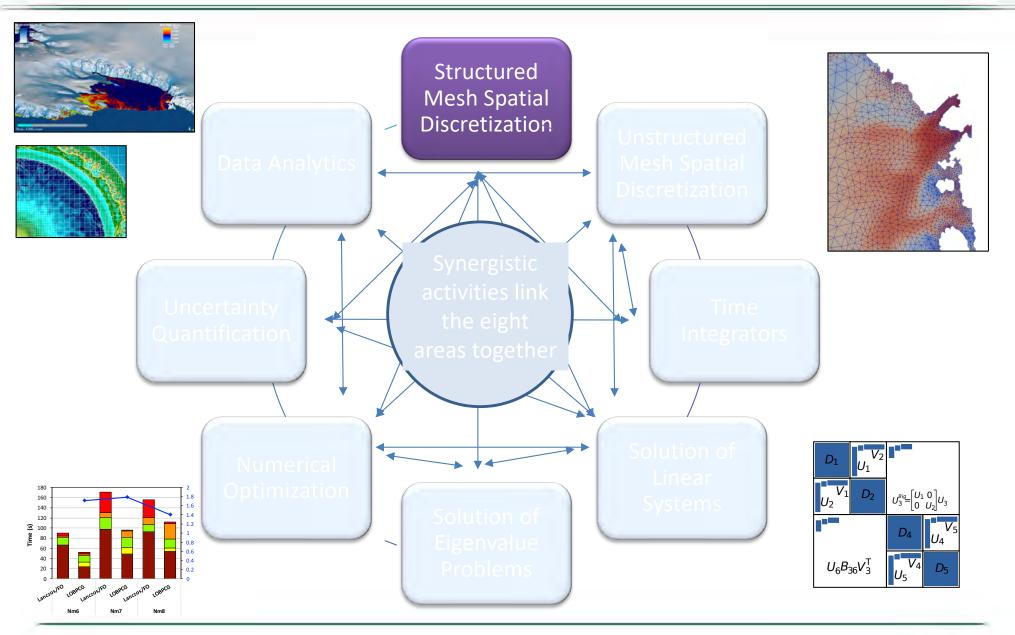
FASTMath is focused on eight core technology areas







Eight core technology areas: Structured Mesh



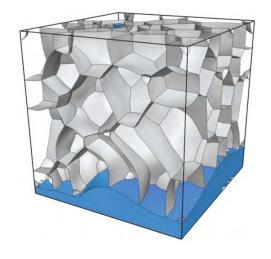


FASTMath Structured Mesh Activities

FASTMath supports several structured mesh frameworks with a wide range of capabilities. Features include support for

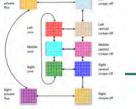
- Adaptive Mesh Refinement (AMR)
- Higher-order interior spatial discretizations
- Higher-order time-stepping
- Higher-order tools for interface dynamics
- Particle dynamics and particle-mesh operations
- Mapped multi-block domains
- Dynamic load balancing
- Interoperability with other solvers, e.g. those in SUNDIALS, PETSc,

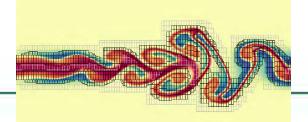












Structured Mesh Application Areas

FASTMath structured grid capabilities are used in numerous DOE projects:

SciDAC-4:

- ComPASS (PIC with AMR)
- TEAMS (astrophysics)
- ProSPect (BISICLES ice sheet model)



- 3D printing
- industrial spray painting
- combustion with electric fields



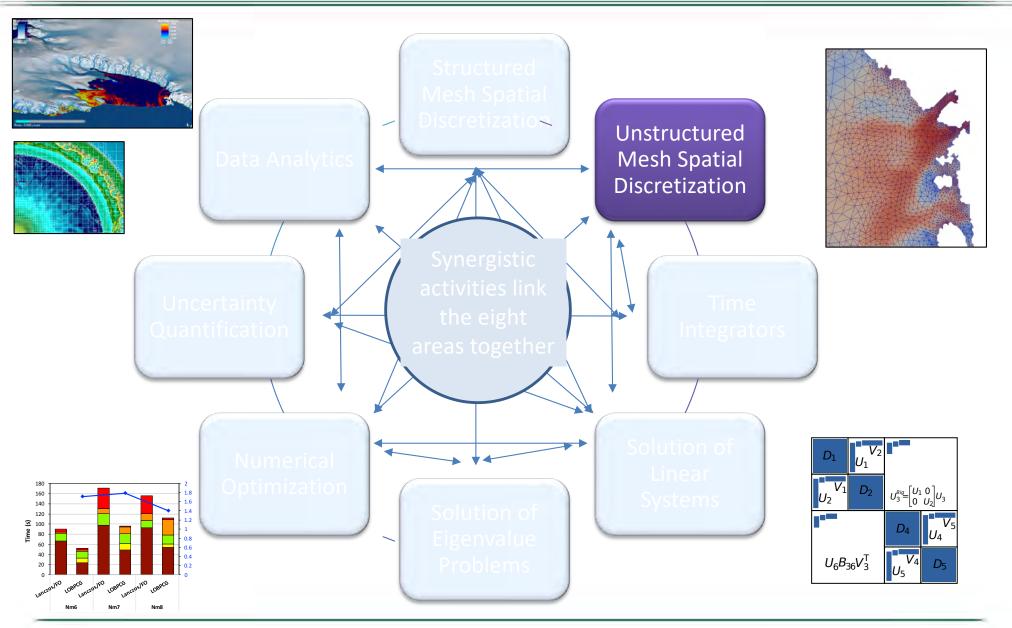


cosmology, multiphase flow, subsurface





Eight core technology areas: Unstructured Meshes

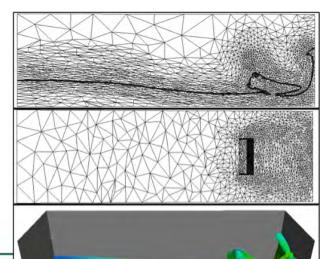




FASTMath Unstructured Mesh Development Areas

FASTMath is providing application developers with tools and methods so they can take direct advantage of the advantages of unstructured mesh technologies.

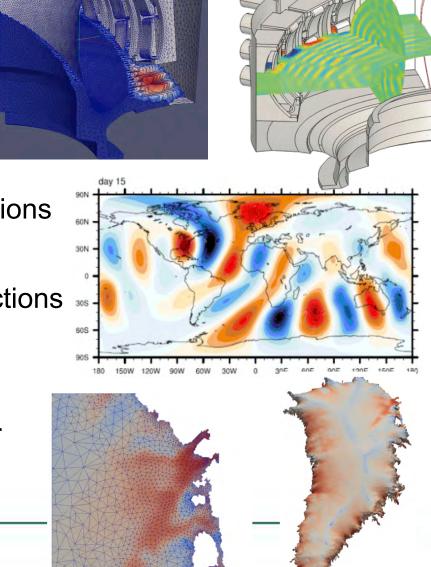
- Unstructured Mesh Analysis Codes
- Unstructured Mesh Adaptation
- Performant Unstructured Meshes
- Dynamic Load Balancing and Task
 Mapping for Extreme Scale Applications
- Unstructured meshes for:
 - PIC methods
 - Management of UQ processes
 - In situ visualization and data analytic





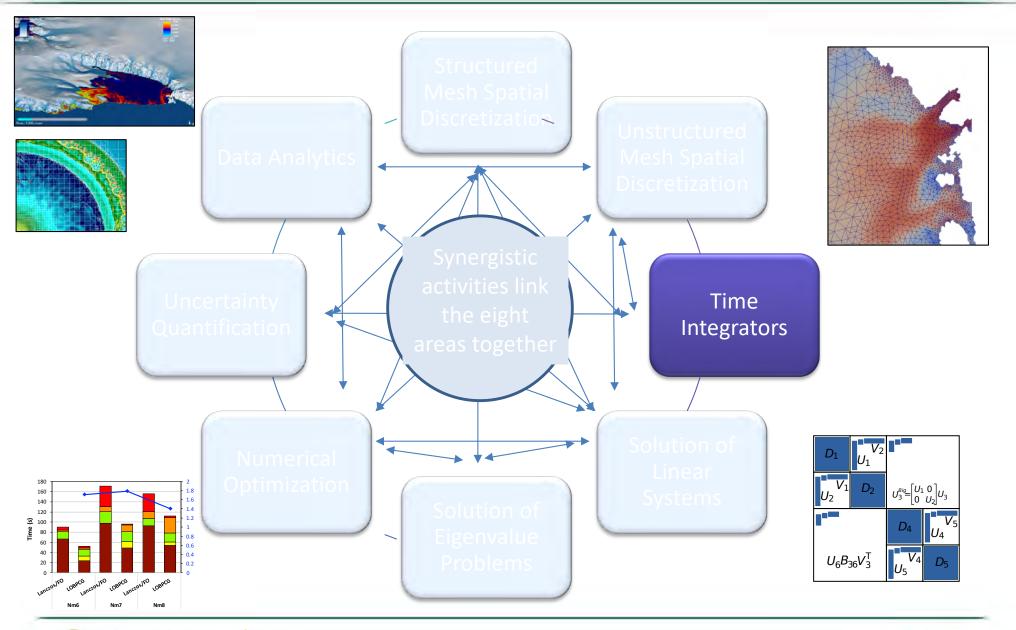
Subset of Current Unstructured Mesh Application Areas

- Analysis of RF waves in tomamak fusion reactors
 - MFEM high-order finite elements
 - PUMI geometry and mesh control
- Atmospheric Modeling
 - Zoltan2's task placement to reduce communications in HOMME's simulations
- Modeling ice sheet melting
 - Core simulation in for sea level predictions
 - FELIX analysis code
 - Omega_h mesh adaptation
- Unstructured mesh methods for 4 fusion SciDACS





Eight core technology areas: Time Integrators





FASTMath activities support several time integration methodologies and software

FASTMath supports:

- Multistep and multistage adaptive time stepping software for ordinary differential equations and differential algebraic equations
- Spectral deferred correction methods and software that include iterative and parallel methods for the time domain
- Adjoint integration methods and software (for discrete and continuous adjoints) for application to optimization contexts

Planned activities:

- Multirate methods in SUNDIALS
- Parallel-in-time methods in SUNDIALS
- High order SDC methods in AMReX
- Multilevel and implicit/explicit SDC methods in AMReX
- Second-order discrete adjoint capabilities in PETSc



FASTMath time integration methods are delivered through software and personnel expertise

Software

- SUNDIALS (https://computation.llnl.gov/projects/sundials)
- PETSc (<u>www.mcs.anl.gov/petsc</u>)
- SDC in AMReX (amrex-codes.github.io)
- Libpfasst (pfasst.lbl.gov)

Personnel:

LLNL: Woodward, Gardner, Loffeld

LBNL: Minion

ANL: Zhang, Smith

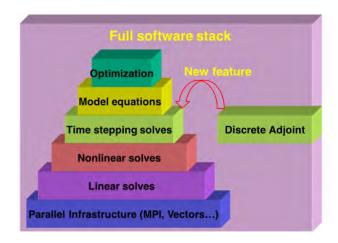
SMU: Reynolds

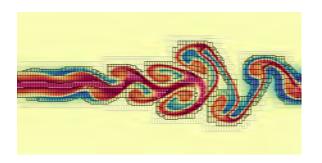


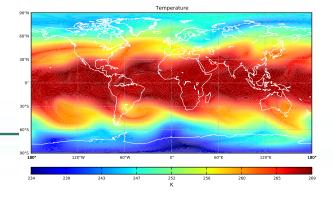
FASTMath activities target innovation in methods and software that will impact applications

DOE applications targeted:

- Climate
 - atmospheric dynamics and physics
 - ice sheets
- Combustion
- Power grid
- Cosmology
- Subsurface flow
- Additive manufacturing

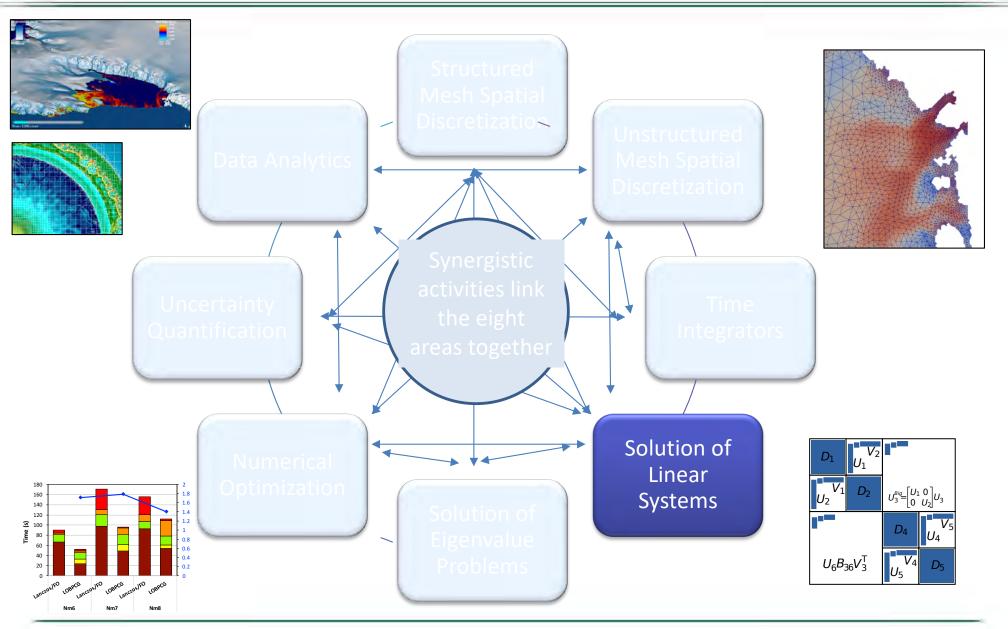








Eight core technology areas: Linear Systems

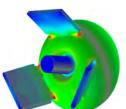




Office

FASTMath Linear Solvers Impact Many Applications

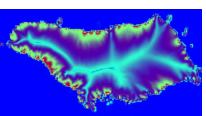
 FASTMath provides a variety of linear solvers to address different needs of DOE applications, such as climate, astrophysics, nuclear physics, fusion, subsurface flow, additive manufacturing, power grid



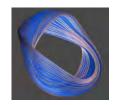
Electromagnetics

- Tokamak Transient Simulations
- Land ice simulations

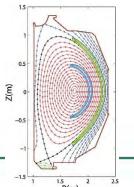
 Simulation of Energetic Particles in Burning Plasmas

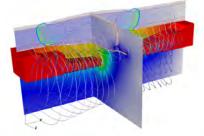


ITER Tokamak



Magnetohydrodynamics



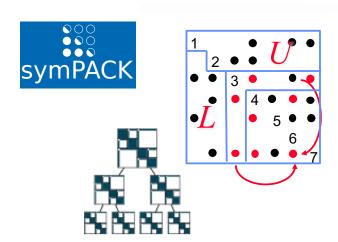


Magnetic flux compression generator simulation enabled by MG smoother research



FASTMath Linear Solvers Software

- SuperLU LU factorizations
- symPACK symmetric positive definite matrices
- STRUMPACK hierarchical basis preconditioners

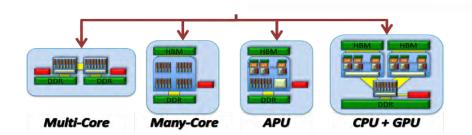


- PDSLin parallel domain decomposition Schur complement based linear solver
- Local discrete convolution methods (LDCMs) for constant coefficient PDEs
- ShyLU (Trilinos) domain decomposition and node level solvers, including BDDC and GDSW preconditioners, Multithreaded LU, Cholesky, and triangular solvers



FASTMath Linear Solvers Software

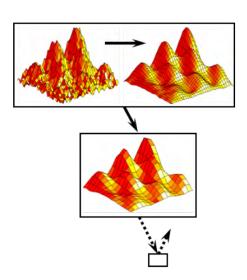
 KokkosKernels: basic linear algebra kernels



 Highly scalable multigrid solvers and preconditioners with different flavors

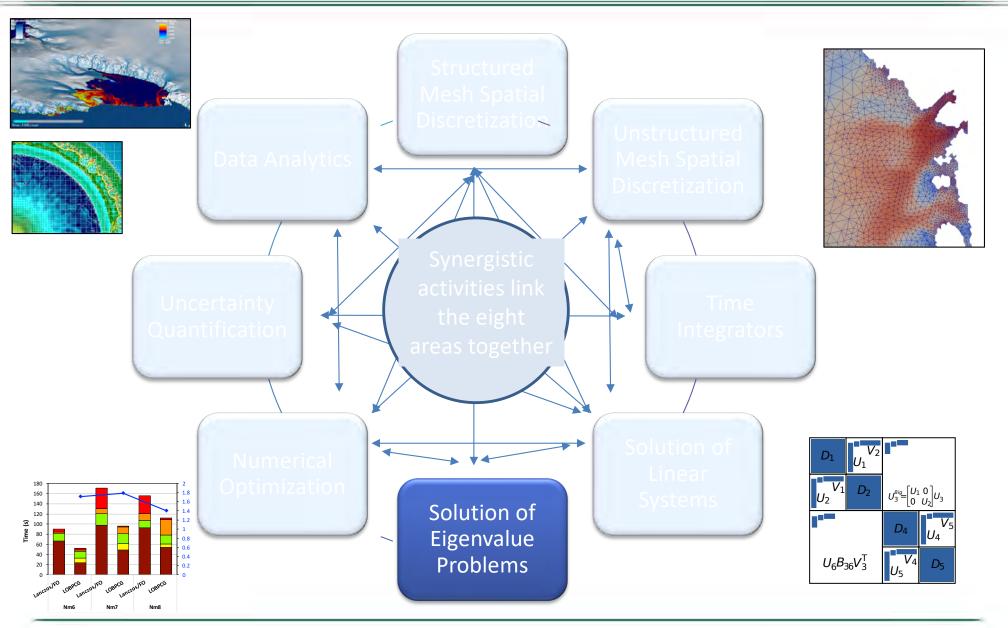
hype high performance preconditioners

- Hypre: structured and unstructured multigrid solvers, conceptual interfaces
- MueLU (Trilinos): smoothed aggregation multigrid solvers
- GAMG (PETSc): geometric, algebraic, hybrid options
- Krylov solvers (PETSc, hypre, ...)





Eight core technology areas: Eigenvalue Problems





Office o Science

FASTMath Eigensolver Activities

FASTMath supports numerical solutions of large-scale eigenvalue problems of a variety types.

- Hermitian eigenvalue problems
- Non-Hermitian eigenvalue problems
- Nonlinear eigenvalue problems with eigenvector nonlinearity

$$H(X)X = X\Lambda$$

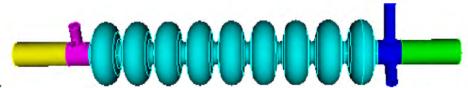


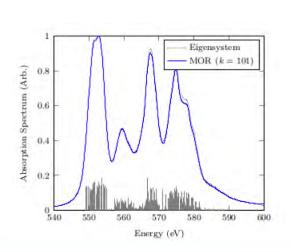
$$T(\lambda)x = 0$$

Eigenvalue problems with Tensor structures

$$H = \sum_{i=1}^{L} I \otimes \cdots \otimes A_i \otimes A_{i+1} \otimes I \cdots$$

Linear response eigenvalue problem







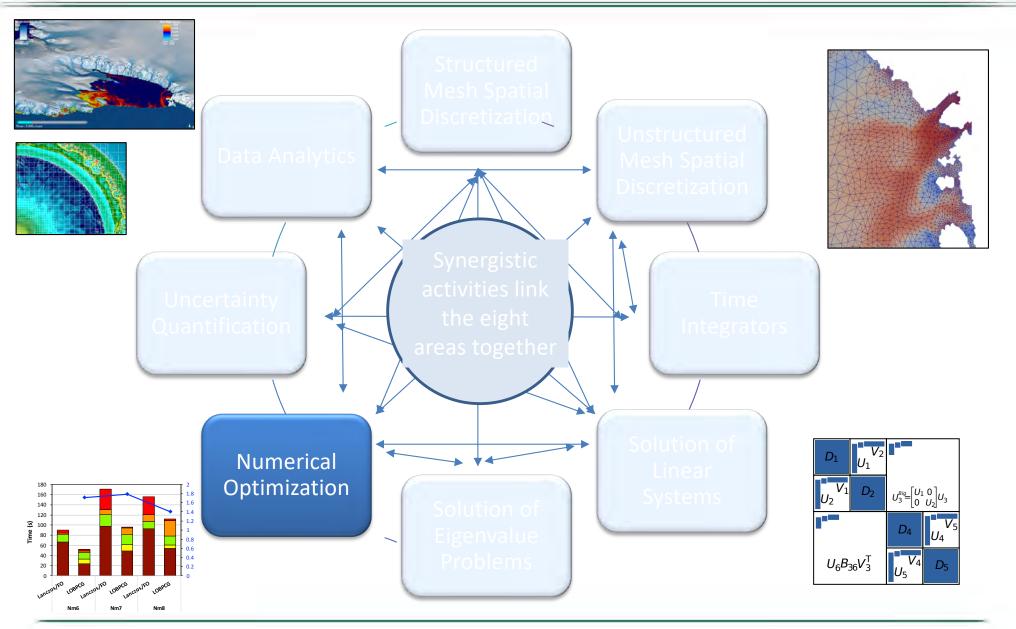
Eigensolver Application areas

FASTMath eigensolvers are used in several DOE applications

- SciDAC-4
 - Catalysis (Kohn-Sham DFT nonlinear eigenvalue problem)
 - Topological materials (Tensor eigenvalue problem)
 - Nuclei structure (Many-body eigenvalue problem)
- Computational Materials Center (linear response eigenvalue problem, spectroscopy)
- Computational Chemistry Center (linear response eigenvalue problem, spectroscopy)
- ECP (NWChem)



Eight core technology areas: Numerical Optimization





For more information contact: Todd Munson (ANL), tmunson@mcs.anl.gov

Numerical Optimization Activities

Develop methods for numerical optimization problems with constraints and for sensitivity analysis using adjoint capabilities.

Dynamic Optimization

- Deliver new capabilities for problems with PDE constraints that include:
 - Dynamics and controls
 - State and design constraints
 - Discrete variables
 - Multiple objectives
- Support a range of derivative requirements

Sensitivity Analysis

- Develop iterative sampling methods that employ sensitivity analysis and surrogate models to determine the most important parameters
- Explore multilevel approach that uses a low-fidelity model to predict parameter sensitivities

Adjoints

- Develop advanced adjoint and forward sensitivity capabilities to provide derivative information
- Provide methods for computing the action of second-order adjoints
- Support calculations involving several quantities of interest

For more information: Todd Munson (tmunson@mcs.anl.gov)



Numerical Optimization Overview

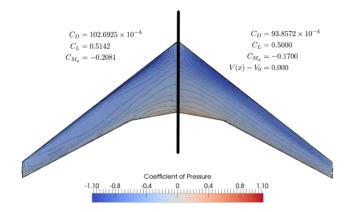
minimize $\mathcal{J}(\mathbf{x}, \mathbf{u})$ x,u

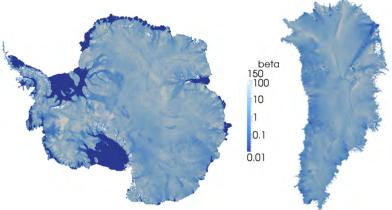
subject to $\mathbf{C}(\mathbf{x}, \mathbf{u}) \leq \mathbf{0}$

 $\mathbf{R}(\mathbf{x}, \mathbf{u}) = \mathbf{0}$

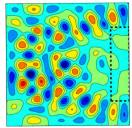
x : Design variables

u: State variables









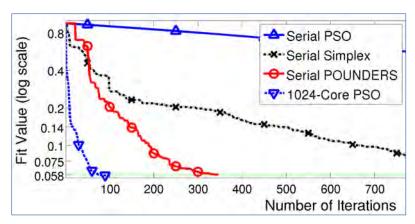
Left: aerodynamic shape optimization Middle: design of electromagnetic scatterer

Right: ice sheet inversion

Numerical Optimization Software and Applications

Software Development

- APOSMM, MATSuMoTo, ORBIT, POUNDERS
 - Derivative-free optimization
- MINOTAUR
 - Mixed-integer nonlinear optimization •
- TAO, ROL
 - PDE-constrained optimization

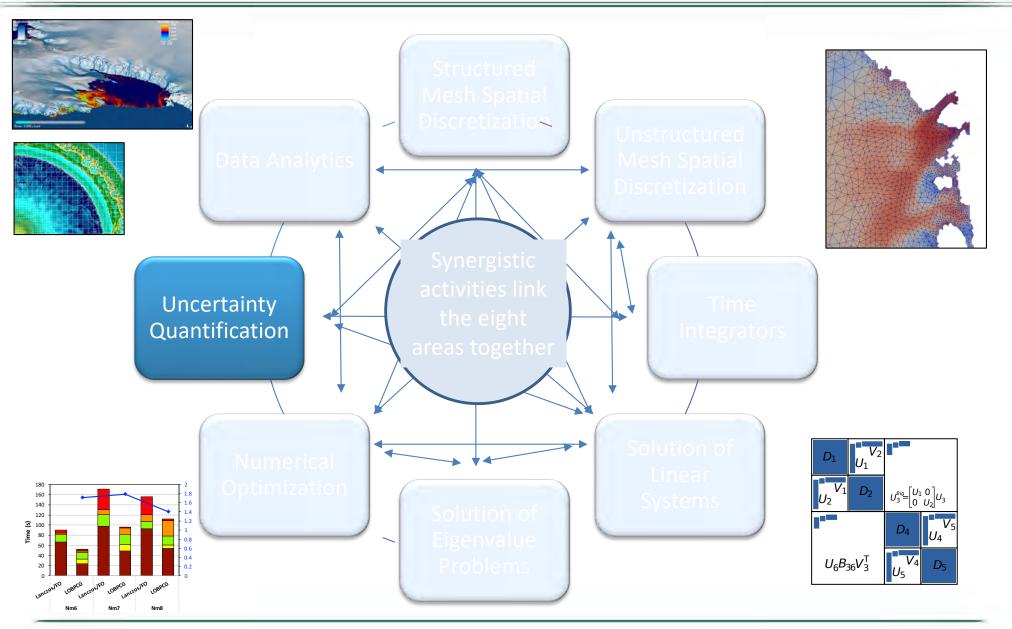


Exploiting structure in particle accelerator calibration

Applications

- NP: Nuclear Computational Low Energy Initiative (NUCLEI)
 - Calibration and optimization of energy density functionals
 - HEP: Community Project for Accelerator Science & Simulation (ComPASS4)
 - Particle accelerator design
- HEP: Accelerating HEP Science: Inference and Machine Learning at Extreme Scales
 - Multi-fidelity optimization and Bayesian parameter estimation
- HEP: Data Analytics on HPC
 - Least-squares problems with integer variables and sensitivity analysis
- BER: Probabilistic Sea-Level Projections from Ice Sheet and Earth System
 - Transient PDE-constrained optimization

Eight core technology areas: Uncertainty Quantification

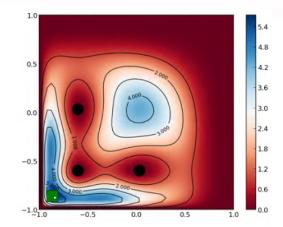


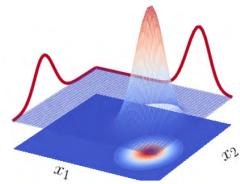


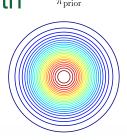
Uncertainty Quantification – Ongoing work

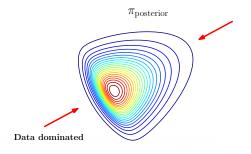
High dimensional function representations

- Polynomial regression, compressive sensing
- Low rank tensors, basis adaptation, manifolds
- Multilevel Multifidelity (MLMF)
 - Monte Carlo, PCE via sparse/low-rank
 - Resource allocation across model form and discretization level hierarchies
- Bayesian inference
 - Data-informed subspace, model error
- Optimization under uncertainty (OUU)
 - Recursive trust region model management with deep hierarchies
 - Reliability-based OUU based on efficient estimation of rare events











Description of the UQ software tools

DAKOTA

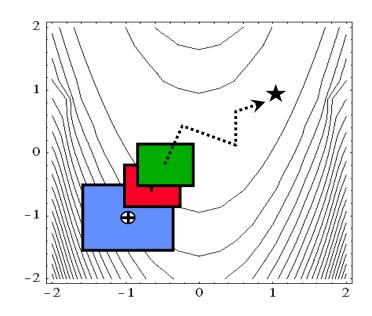
- Uncertainty quantification and Optimization
- Open source (GNU LGPL)
- dakota.sandia.gov

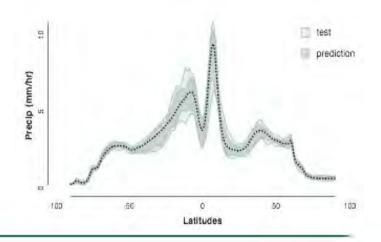
UQ Toolkit – UQTk

- Uncertainty quantification
- Open source (GNU LGPL)
- www.sandia.gov/UQToolkit





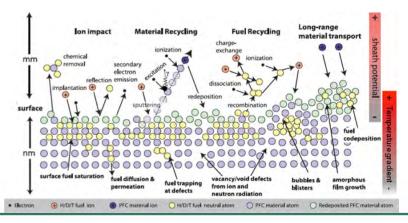


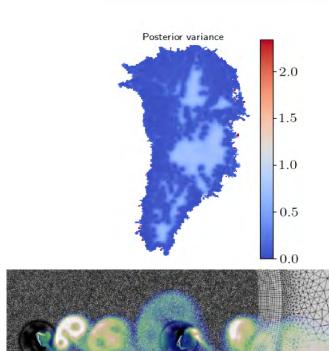


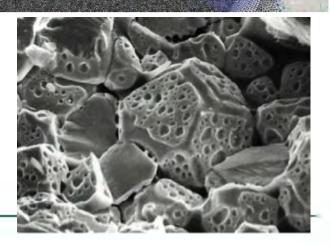


UQ Application interactions

- Ice sheet and earth system modeling
- Tokamak disruption simulation
- Exascale wind flow modeling
- E3SM climate modeling
- Sensor networks for climate modeling
- Fusion plasma surface interactions
- Fission gas in nuclear fuel

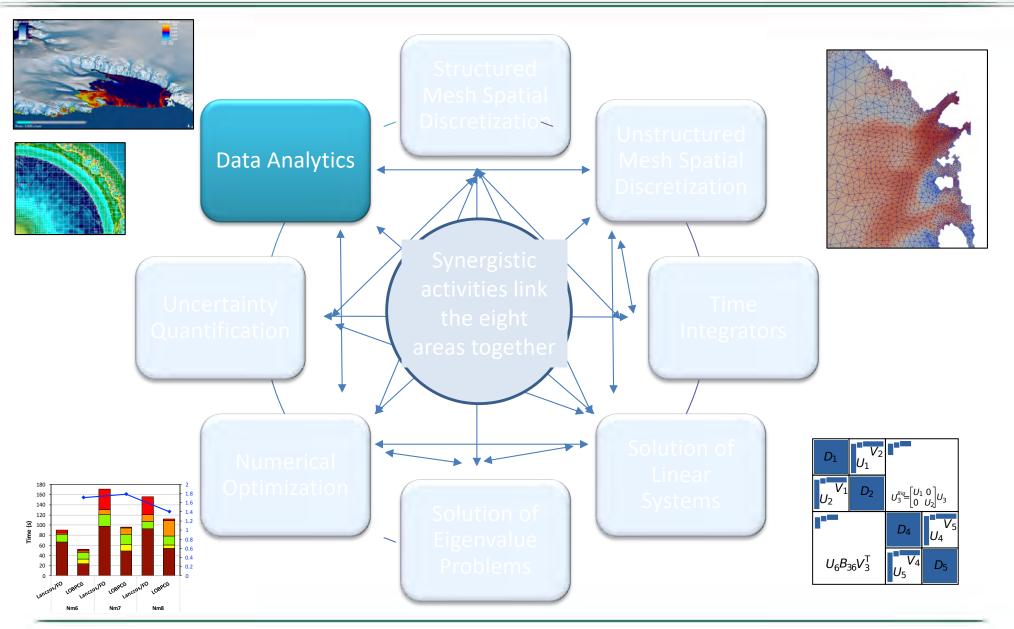








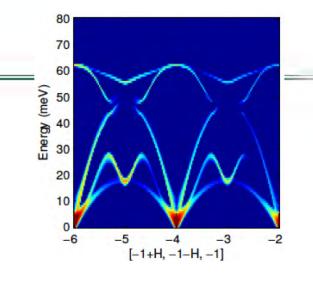
Eight core technology areas: Data Analytics





Data Analytics

- Goal: Sparse functional representation of data, to enable faster IO and analysis of big datasets
- Software tools: Tasmanian, PUMI, TAO
- FASTMath Tasks



Sparse IO of Big Data

- Advanced sampling techniques for distributed data
- Research adaptive methods for sparse representation
- Build accurate uncertainty estimates for sparse representations of data

Fast Estimation & Evaluation

- Develop and design high order regularizes that optimize functional representations of data
- Surrogate models that accelerate estimation and evaluation of sparse data approximation

Streaming Data Ranking

- Algorithms that maximize information transfer
- Ordered sparse functional representations of data
- Parallel methods for streaming distributed datasets

For more information: Rick Archibald (<u>ArchibaldRK@ornl.gov</u>), Clayton Webster (ORNL), & Hoang Tran (ORNL)



Data Analytics: Technical description of technology

We reconstruct data $c \in \mathbb{R}^{N \times l}$ from measurements $u \in \mathbb{R}^{m \times l}$ and $A \in \mathbb{R}^{m \times N}$:

$$u \approx Ac$$

- Limited number of measurements: $m \ll N$.
- The data are sparse.
- l = 1: reconstructing a single dataset.

l > 1: simultaneously reconstructing multiple datasets.

Recovery via regularizations enforcing sparsity:

$$c = \operatorname{argmin} R(z)$$
 subject to $u \approx Az$

Standard CS: $R(z) = ||z||_1$.

Structures of the sparsity can be exploited:

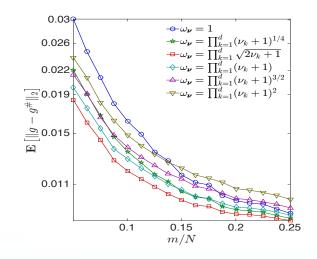
- Downward closed and tree structures: $R(z) = ||z||_{\omega,1}$.
- Joint sparsity: $R(z) = ||z||_{2,1}.$

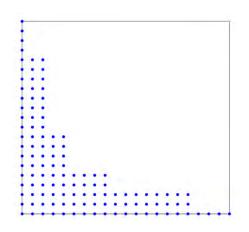
Data Analytics: Technical description of technology

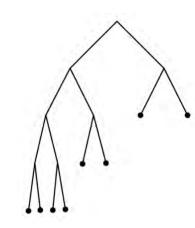
- Data from UQ and imaging applications often possess downward closed and tree structure.
- \blacktriangleright Weighted l_1 minimization with a specific choice of weight:

$$R(z) = ||z||_{\omega,1}$$
 with $\omega_j = \max |A_{:,j}|$

Figure: A comparison of weighted l_1 minimization with different choices of weights







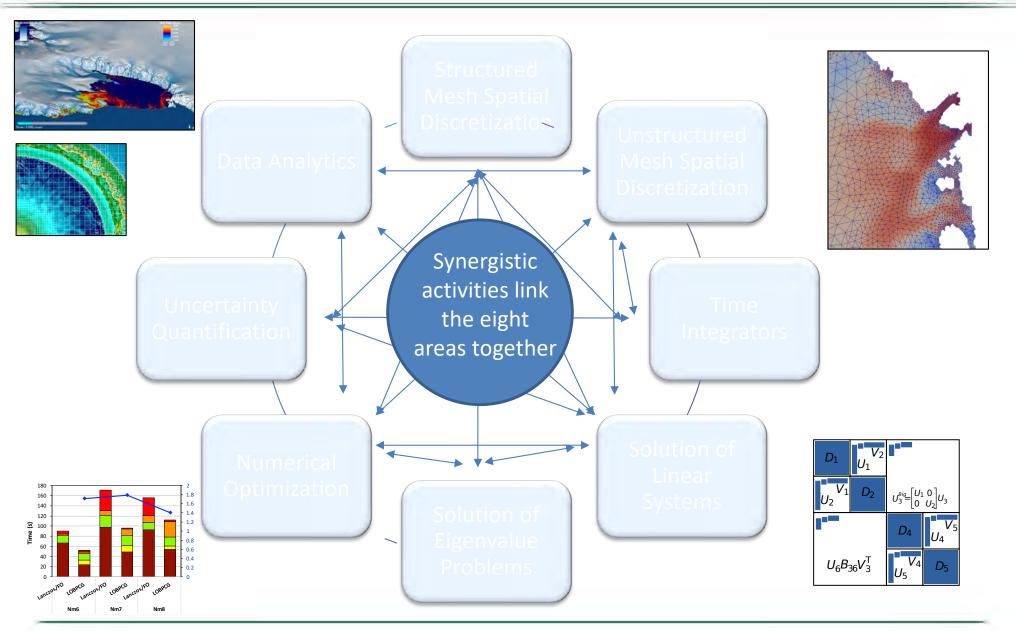
Certified reduction in complexity:

- Legendre systems: $m = O(s^2)$ instead of $O(s^{2.58})$ as in unweighted l_1 .
- Chebyshev systems: $m = O(s^{1.58})$ instead of $O(s^2)$ as in unweighted l_1 .

A. Chkifa, N. Dexter, H. Tran, and C. Webster, *Polynomial approximation via compressed sensing of high-dimensional functions on lower sets*. **Math. Comp.** (2017) https://doi.org/10.1090/mcom/3272



Eight core technology areas: Synergistic Activites







Synergystic activities will results in new capabilities or higher efficiencies

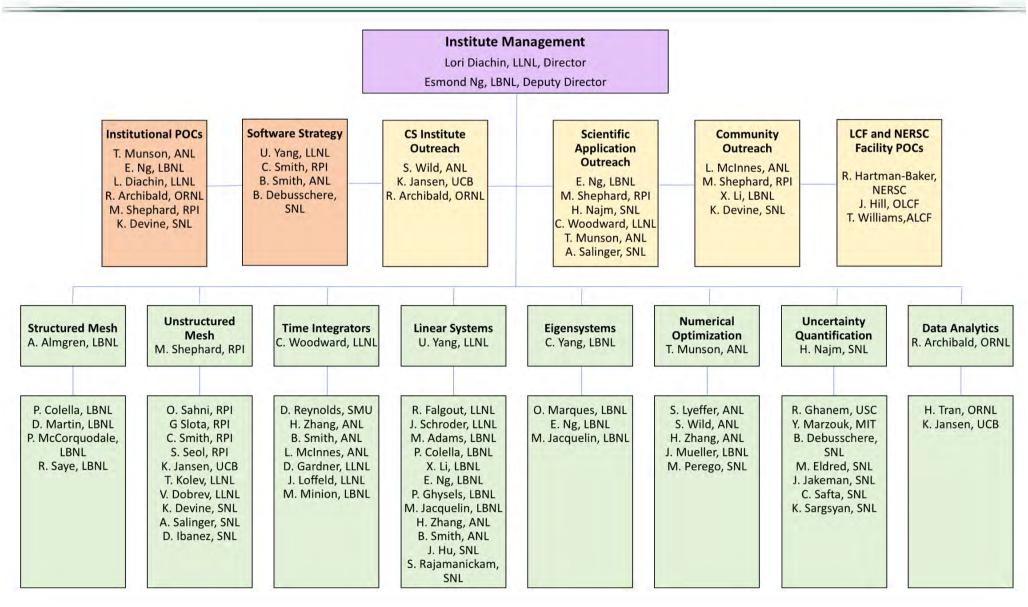
New Capabilities

- Optimization under uncertainty
- Discrete and multi-objective optimization for data analytics
- In situ simulation on unstructured meshes

Higher Efficiency

- Leverage multiple right hand sides from optimization and UQ ensembles in linear and nonlinear solvers
- Adaptivity in the spatial and stochastic space in UQ on unstructured grids;
- Dynamic UQ load balancing
- In situ simulation on unstructured meshes

The FASTMath organizational structure follows from the core technical areas



For more information...

Contacts:

- Lori Diachin, Director (<u>diachin2@Ilnl.gov</u>)
- Esmond Ng, Deputy Director (egng@lbl.gov)
- Any of the core components leads

Web site:

www.fastmath-scidac.org







LLNL Auspices and Disclaimer

Lawrence Livermore National Laboratory is operated by Lawrence Livermore National Security, LLC, for the U.S. Department of Energy, National Nuclear Security Administration under Contract DE-AC52-07NA27344. This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

